

course introduction

CS 585, Fall 2019

Introduction to Natural Language Processing

<http://people.cs.umass.edu/~miyyer/cs585/>

Mohit Iyyer

College of Information and Computer Sciences

University of Massachusetts Amherst

natural language processing

natural language processing

languages that evolved naturally through human use
e.g., Spanish, English, Arabic, Hindi, etc.

NOT: controlled languages (e.g., Klingon)

NOT: programming languages

natural language processing

supervised learning: *map text to **X***

unsupervised learning: *learn **X** from text*

generate text from **X**

how?

- Math! Algorithms!
- Data!
- Code!
 - Skill: translating from math to code
 - Skill: debugging math/linguistic/algorithm code
- A little bit of linguistics goes a long way

who?

TAs:

Tu Vu

Shufan Wang

Simeng Sun

Varun Sharma

email all of us (including me!) at
cs585nlp@gmail.com

course website:

<https://people.cs.umass.edu/~miyyer/cs585>

waitlist override pass/fail etc.

- don't email us about getting into the class because we can't help... please contact Darlene Fahey at fahey@cs.umass.edu with such questions or requests
- anyone can sit in the class!
- for MS students: the pass/fail deadline will be Oct. 29, the same as it is for undergrads

office hours every day of the week!

Monday w/ Shufan: 11:30am-12:30pm in CS207 Cube 3

Tuesday w/ Mohit: 4-5PM in CS258

Wednesday w/ Varun: 12-1PM in CS207 Cube 4

Thursday w/ Simeng: 11am-12pm in CS207 Cube 4

Friday w/ Tu: 3-4PM in CS207 Cube 4

If necessary, TA office hours will be extended by one hour during homework / exam weeks

anonymous questions / comments?

- submit questions/concerns/feedback to <https://forms.gle/j9ECQXX9pJFb4zvg7>
- we will go over some/all submitted responses at the start of every class

Prerequisites

- Comfort with programming, algorithmic thinking
 - Ever debugged a graph algorithm? Know its Big-O time and space requirements?
 - CS 220 or 230
- Comfort with probability and mathematical notation
 - Ever used Bayes Rule?
 - CS 240
- Excitement about language!
- Willingness to learn

Requirements

- (10%) Participation and short exercises
 - Bring pencils/pens/paper to class
- (30%) Problem sets
 - Written: math and concepts
 - Programs: in Python
 - All HWs will be on Google Colab other than HW0
- (25%) Midterm (in class, end of October)
- (35%) Final projects (groups of 5)
 - Choose a topic, or select a suggested topic
 - Project proposal
 - Progress report
 - Final report / presentation

Logistics

- Main course website:
<http://people.cs.umass.edu/~miyyer/cs585/>
- Gradescope for homework submissions
- Moodle for lecture video recordings
- 585-01 and 585-02 sections are the same
- Due next Thursday: HW0

Readings

- No need to buy any textbooks!
- Readings will be provided as PDFs on website
- Often draft chapters from Jurafsky and Martin, *Speech and Language Processing*, or Jacob Eisenstein's *Natural Language Processing*, or random research papers / notes :)

Levels of linguistic structure

Discourse

Semantics

Syntax: Constituents

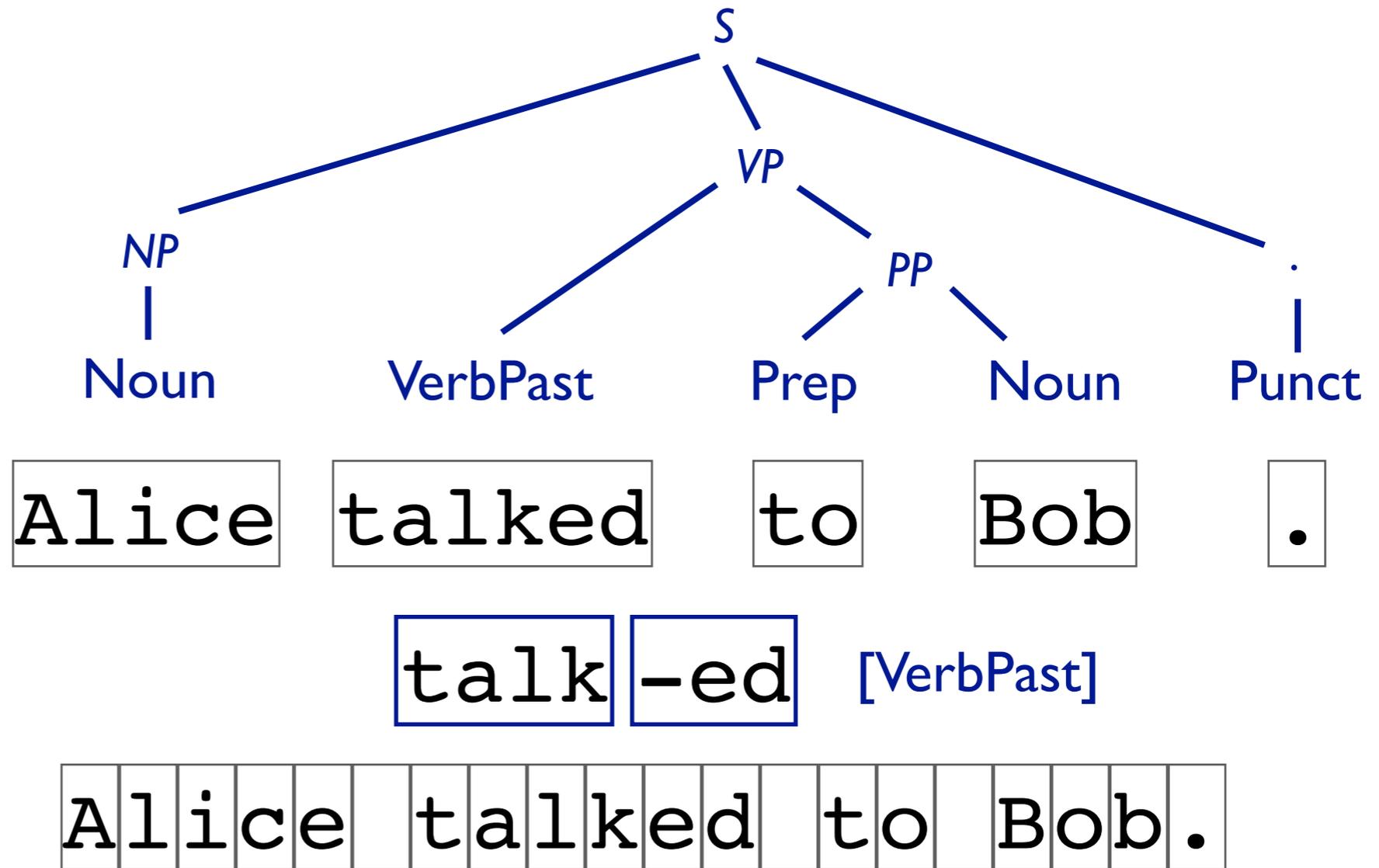
Syntax: Part of Speech

Words

Morphology

Characters

CommunicationEvent(e) SpeakerContext(s)
 Agent(e, Alice) TemporalBefore(e, s)
 Recipient(e, Bob)



demos!
(allennlp.org)

demos!

(<https://talktotransformer.com/>)

python demo!
(colab.research.google.com)

- Check out HW0 on the website
- See you on Thursday